

ue Jun 1 09:38:04 2004

us-10-063-688-34.0196.rspc

Appendix B

Page -

SEQUENCE 678 AA; 73930 MM; 9870E75A218C686C CRC64;

Query Match 100.0%; Score 678; DB 4; Length 678;
 est Local Similarity 100.0%; Pred. No. 0;
 atches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MNTVUIMKASVIEKFLVLTGTHSKNKTAKIKRKTTPQINCVDYAKIIDEFFIV 60
 1 MNTVUIMKASVIEKFLVLTGTHSKNKTAKIKRKTTPQINCVDYAKIIDEFFIV 60
 1 MNTVUIMKASVIEKFLVLTGTHSKNKTAKIKRKTTPQINCVDYAKIIDEFFIV 60
 61 KCPAGGADPKHNVGTVDYASVSCAAVHSGVLDNNGKILVAKVAGSGYKSYNG 120
 61 KCPAGGADPKHNVGTVDYASVSCAAVHSGVLDNNGKILVAKVAGSGYKSYNG 120
 121 VQSLSPRRESFVLESKPKKGVTPSALTYSKSPAAQGEFTKAYORPPIPTTA 180
 121 VQSLSPRRESFVLESKPKKGVTPSALTYSKSPAAQGEFTKAYORPPIPTTA 180
 121 VQSLSPRRESFVLESKPKKGVTPSALTYSKSPAAQGEFTKAYORPPIPTTA 180
 181 PVTLMQILAVYAVATPTLPRPSPASSTSIPEOSVGRSOMQIMSTATTSSOR 240
 181 PVTLMQILAVYAVATPTLPRPSPASSTSIPEOSVGRSOMQIMSTATTSSOR 240
 181 PVTLMQILAVYAVATPTLPRPSPASSTSIPEOSVGRSOMQIMSTATTSSOR 240
 241 PRADPGIOPDPSGAAPVGVADVSGLVPEKELSTGSEFVSIGDPNCKIDLSFLID 300
 241 PRADPGIOPDPSGAAPVGVADVSGLVPEKELSTGSEFVSIGDPNCKIDLSFLID 300
 241 PRADPGIOPDPSGAAPVGVADVSGLVPEKELSTGSEFVSIGDPNCKIDLSFLID 300
 301 STSIGRRFRIOKQILADVAQALDIPAGPLMGVQYGDNPATPHNLKTHNSRDLKTAI 360
 301 STSIGRRFRIOKQILADVAQALDIPAGPLMGVQYGDNPATPHNLKTHNSRDLKTAI 360
 301 STSIGRRFRIOKQILADVAQALDIPAGPLMGVQYGDNPATPHNLKTHNSRDLKTAI 360
 361 EKTTOGGLSNVGRALISFVTKNPFSSKANGRSAPVNVVMDGPTDKVEASRLAES 420
 361 EKTTOGGLSNVGRALISFVTKNPFSSKANGRSAPVNVVMDGPTDKVEASRLAES 420
 361 EKTTOGGLSNVGRALISFVTKNPFSSKANGRSAPVNVVMDGPTDKVEASRLAES 420
 421 GINIFETITTEGAENKQYVEEPNPAKAYCPNNGFYSILHOSMGKAKTLOIYKVC 480
 421 GINIFETITTEGAENKQYVEEPNPAKAYCPNNGFYSILHOSMGKAKTLOIYKVC 480
 421 GINIFETITTEGAENKQYVEEPNPAKAYCPNNGFYSILHOSMGKAKTLOIYKVC 480
 481 TDRLASCTCLNSADIGVLDGSSVGTGNFVLOFTNLTKEFISDTDIRIGAVQT 540
 481 TDRLASCTCLNSADIGVLDGSSVGTGNFVLOFTNLTKEFISDTDIRIGAVQT 540
 481 TDRLASCTCLNSADIGVLDGSSVGTGNFVLOFTNLTKEFISDTDIRIGAVQT 540
 541 YEOBLEFPGDKYSXPDIINAIRKVGWVSGSTGAINFALEOLFKSKENKRMKMLI 600
 541 YEOBLEFPGDKYSXPDIINAIRKVGWVSGSTGAINFALEOLFKSKENKRMKMLI 600
 541 YEOBLEFPGDKYSXPDIINAIRKVGWVSGSTGAINFALEOLFKSKENKRMKMLI 600
 601 TDRSSTDVRIIPAMAHLKGVITTAIGVANAQBELVATAPARHSPFFVDEFDNLHOY 660
 601 TDRSSTDVRIIPAMAHLKGVITTAIGVANAQBELVATAPARHSPFFVDEFDNLHOY 660
 601 TDRSSTDVRIIPAMAHLKGVITTAIGVANAQBELVATAPARHSPFFVDEFDNLHOY 660
 661 VPRIONICTEFGNSOPRN 678
 661 VPRIONICTEFGNSOPRN 678
 661 VPRIONICTEFGNSOPRN 678

RESULT 2

Q96DM8 PRELIMINARY; PRT; 693 AA.

ID Q96DM8
 AC Q96DM8
 DT 01-DEC-2001 (T-EMBLrel. 19, Created)
 DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Hypothetical protein FLJ32210.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RA Ishibashi T., Kanohori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
 RA Horiuchi T., Hirooka S., Murakawa K., Takiguchi S., Kusano J.,
 RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,

RA Sugiyama T., Irie R., Otsubi T., Sato H., Wakamatsu A., Ishii S.,
 RA Yamamoto J., Isono Y., Kawata-Hio Y., Saito K., Nishikawa T.,
 RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
 RA Nagatsuma M., Takahashi-Pujil A., Oshima A., Sugiyama A., Kawakami B.,
 RA Suzuki Y., Sugano S., Nagahara K., Masuno Y., Nagai K., Itoigai T.,
 RT "HBD human CDNA sequencing project."
 RL Submitted (Oct-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AK056772; BAB1279.1; -
 DR Genbank; HGNC:12697; VIT.
 DR InterPro; IPR004043; LCCU_dom.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF03815; LCCU; 1.
 DR Pfam; PRO0092; vwa; 2.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00603; LCCU; 1.
 DR SMART; SM00327; VWA; 2.
 DR PROSITE; PS00820; LCCU; 1.
 DR PROSITE; PS00334; VWF; 2.
 DR Hypothetical protein
 SQ SEQUENCE 693 AA; 75575 MM; 2DE8B2421F2D496D CRC64;

Query Match 60.6%; Score 411; DB 4; Length 693;
 est Local Similarity 100.0%; Pred. No. 0;
 Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 268 GLVPEKELSTGSEFVSIGDPNCKIDLSFLIDGSTRIGKARFRIQKQILADVAQALDIP 327
 DB 283 GLVPEKELSTGSEFVSIGDPNCKIDLSFLIDGSTRIGKARFRIQKQILADVAQALDIP 342
 QY 328 AGPLMGVQYGDNPATPHNLKTHNSRDLKTAIEKITOGGLSNVGRALISFVTKNPFSSKA 387
 DB 343 AGPLMGVQYGDNPATPHNLKTHNSRDLKTAIEKITOGGLSNVGRALISFVTKNPFSSKA 402
 QY 388 NGNRSAPVNVVMDGPTDKVEASRLAESGINIFETITTEGAENKQYVEEPNPAK 447
 DB 403 NGNRSAPVNVVMDGPTDKVEASRLAESGINIFETITTEGAENKQYVEEPNPAK 462
 QY 448 KAVCTNGFISLHVSQSPFGAKTLOIYKVCVDTDRLASCTCLNSADIGVLDGSSVIG 507
 DB 463 KAVCTNGFISLHVSQSPFGAKTLOIYKVCVDTDRLASCTCLNSADIGVLDGSSVIG 522
 QY 508 TGNFETVLOFTNLTKEFISDTDIRIGAVQTYEORLEFPGDKYSXPDIINAIRKVG 567
 DB 523 TGNFETVLOFTNLTKEFISDTDIRIGAVQTYEORLEFPGDKYSXPDIINAIRKVG 582
 QY 568 WSGGSTGAALNFALEOLFKSKENKRMKMLITDGRSYDVRIPAMAHLKGVITTAIG 627
 DB 589 WSGGSTGAALNFALEOLFKSKENKRMKMLITDGRSYDVRIPAMAHLKGVITTAIG 642
 QY 628 VAMAQBELVATAPARHSPFFVDEFDNLHOYVRIQNICTEFGNSOPRN 678
 DB 643 VAMAQBELVATAPARHSPFFVDEFDNLHOYVRIQNICTEFGNSOPRN 693

RESULT 3

Q96DT1 PRELIMINARY; PRT; 656 AA.

ID Q96DT1
 AC Q96DT1
 DT 01-DEC-2001 (T-EMBLrel. 19, Created)
 DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Vitrin.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle, and fetal heart;
 RA Ren Z.-X., Liu J.G., Mayne R.,
 RA "Human vitrin complete cDNA sequence."
 RT Submitted (May-1998) to the EMBL/Genbank/DBJ databases.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2004, 16:47:44 ; Search time 52 Seconds
(without alignments)
413.872 Million cell updates/sec

Title: US-10-063-688-34

Sequence: 1 MRTLVLTKKASVIEMLVTL.....QVPRIONICTERNSPRN 678

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size: 6

Total number of hits satisfying chosen parameters: 8650

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	678	100.0	678	4	Q9UDNO
2	411	60.6	693	4	Q96DM8
3	259	38.2	656	4	Q96DT1
4	57	8.4	652	6	Q95L12
5	45	6.6	628	11	Q8BQ41
6	45	6.6	650	11	Q8K047
7	45	6.6	650	11	Q8CY21
8	45	6.6	650	11	Q8VH15
9	11	1.6	553	13	Q8AM56
10	8	1.2	100	10	Q8H8Y2
11	8	1.2	112	16	Q8G578
12	8	1.2	112	16	Q8CB89
13	8	1.2	208	13	Q90Z19
14	8	1.2	230	2	Q8KWO1
15	8	1.2	230	2	Q8KWO1
16	8	1.2	231	2	Q86887

17	8	1.2	231	2	Q9RIP2	Q9RIP2 streptococc
18	8	1.2	231	2	Q9AHD3	Q9AHD3 streptococc
19	8	1.2	285	16	Q92M09	Q92M09 rhizobium m
20	8	1.2	359	1	Q47973	Q47973 halobacteri
21	8	1.2	368	17	Q9HPU7	Q9HPU7 halobacteri
22	8	1.2	382	12	Q83907	Q83907 ovine adeno
23	8	1.2	391	16	Q98HP8	Q98HP8 rhizobium l
24	8	1.2	394	16	Q97FE4	Q97FE4 clostridium
25	8	1.2	400	16	Q7UD27	Q7UD27 shigella fl
26	8	1.2	416	5	Q8IRF0	Q8IRF0 drosophila
27	8	1.2	416	16	Q8FJ98	Q8FJ98 escherichia
28	8	1.2	416	16	Q831N3	Q831N3 shigella fl
29	8	1.2	461	3	Q9HRC6	Q9HRC6 neosporea
30	8	1.2	462	2	Q8RMG7	Q8RMG7 acetobacter
31	8	1.2	494	4	Q961U6	Q961U6 homo sapien
32	8	1.2	516	4	Q9H647	Q9H647 homo sapien
33	8	1.2	559	16	Q82114	Q82114 streptomyce
34	8	1.2	573	5	Q9GV07	Q9GV07 dugesia dor
35	8	1.2	686	5	Q9W013	Q9W013 drosophila
36	8	1.2	695	16	Q9CH87	Q9CH87 lactococcus
37	8	1.2	718	6	Q8KXH0	Q8KXH0 macaca fasc
38	8	1.2	763	5	Q960Y7	Q960Y7 drosophila
39	8	1.2	892	16	Q99PX4	Q99PX4 streptomyce
40	8	1.2	1078	5	Q9NDY3	Q9NDY3 leishmania
41	8	1.2	1464	5	Q9VX75	Q9VX75 drosophila
42	8	1.2	1464	5	Q96782	Q96782 drosophila
43	8	1.2	2197	5	Q9W016	Q9W016 drosophila
44	7	1.0	29	11	Q8VIF4	Q8VIF4 mus musculu
45	7	1.0	34	4	Q9H3R8	Q9H3R8 homo sapien

ALIGNMENTS

RESULT 1

ID	Q9UDNO	PRELIMINARY	PRT	678 AA.
AC	Q9UDNO			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	Similar to Coch-SB2.			
GN	WDSC:H.NH029411.1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RM	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99063792; PubMed=9847074;			
RA	Sulston J.E., Waterston R.;			
RT	"Toward a complete human genome sequence.";			
RL	Genome Res. 8:1097-1108(1998).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RT	Cordes M., Kalicki J., Ames M.;			
RA	"The sequence of Homo sapiens BAC clone RP11-29411.1.";			
RL	Submitted (APR-1999) to the EMBL/GenBank/DBD databases.			
RM	[3]			
RP	SEQUENCE FROM N.A.			
RA	Waterston R.;			
RL	Submitted (DEC-1999) to the EMBL/GenBank/DBD databases.			
DR	EMBL; AC007363; AAF19243.1; -.			
DR	HSSP; P11215; IJLM.			
DR	InterPro; IPR004043; LCCL dom.			
DR	InterPro; IPR02035; VWF_A.			
DR	Pfam; PF00092; VWF_2.			
DR	PRINTS; PR00453; VWFADOMAIN.			
DR	SMART; SM00603; LCCL; 1.			
DR	SMART; SM00327; VWA; 2.			
DR	PROSITE; PS50820; LCCL; 1.			
DR	PROSITE; PS50234; VWA; 2.			